DOCUMENT RESUME

ED 215 569 FL 012 863

AUTHOR	Fischer, Kathleen B.; Caldwell, J.
TITLE	Establishing Criteria for Transition of ESL Students to Monolingual English Classrooms.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Aug 78
GRANT	OB-NIE-G-78-0213
NOTE	46p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Bilingual Education; Dominicans; *English (Second Language); Grade 3; Grade 4; Intermediate Grades; *Language Proficiency; *Language Tests; Listening Comprehension; Mexican Americans; Mexicans; Primary Education; Puerto Ricans; Reading Comprehension; Reading Instruction; Regression (Statistics); Spanish; Syntax; Teacher Attitudes; Test Reliability; Test Validity; *Transitional Programs; Vocabulary Skills

ABSTRACT

This study explores several research questions. Of primary concern was the identification of variables that might be employed to predict the potential of students to succeed in monolingual English classrooms following transition from bilingual programs. Although consideration was given to various background and demographic factors, the principal focus of the research was on language proficiency variables--specifically reading comprehension, verbal production, and aural comprehension in English and the native language, Spanish. Of secondary and collateral concern was the validity and reliability of available instruments that might be used to measure these various language proficiencies. Two batteries of tests were constructed and refined. The appendix provides a list of the objectives or skills that were specified for construction of the batteries. A third concern was the comparability of results between different Hispanic language groups--students of predominantly native Mexican and Mexican American origin in the Los Angeles area and students of predominantly Puerto Rican and Dominican origin from the East Coast of the U.S. Analysis explores the relationship between language proficiency and classroom performance prior to and following transition for the current year's third-grade students enrolled in transitional bilingual programs, and the students who had been enrolled in third grade transitional bilingual programs the previous year. (Author/JK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED215569

DELIVERABLE: AUGUST 1978

STUDIES IN BILINGUAL EVALUATION

Work Unit 1:

Bilingual Prediction Study

Kathleen B. Fischer
Project Director

ESTABLISHING CRITERIA FOR TRANSITION OF ESL STUDENTS TO MONOLINGUAL ENGLISH CLASSROOMS

K. B. Fischer and J. Caldwell

Grant Number
OB-NIEG-78-0213

FL012863

Center for the Study of Evaluation
Graduate School of Education
University of California, Los Angeles

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Center for the Study of Evaluation
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ESTABLISHING CRITERIA FOR TRANSITION OF ESL STUDENTS TO MONOLINGUAL ENGLISH CLASSROOMS

In January, 1974, the U.S. Supreme Court ruled, in the case of Lau v. Nichols (414 U.S. 563, 39 L ED 2d 1, 94 Ct 786), that a school district receiving Federal funds must ensure that non-English-speaking students acquire the basic skills necessary to profit from the regular instruction being provided by the district. In the opinion of the Court:

> There is no equality of treatment merely by providing students with the same facilities, textbooks, teachers, and curriculum; for students who do not understand English are effectively foreclosed from any meaningful education... those who do not understand English are certain to find their classroom experience wholly incomprehensible and in no way meaningful.

The Court found the San Francisco Unified School District, therefore, to be in violation of:

a) Section 601 of the 1964 Civil Rights Act which bars discrimination based "on the ground of race, color, or national origin" in "any program or activity receiving Federal financial assistance," and

b) of the implementing HEW regulations that require districts to "take affirmative steps to rectify the language deficiency in order to open its instructional program to these students."

In the summer of 1975, the Department of Health, Education, and Welfare issued recommendations for meeting the requirements of the Lau v. Nichols decision. Included among educational approaches that would constitute appropriate "affirmative steps" at the early elementary grades are bilingual/bicultural education programs and transitional bilingual

education (TBE) programs. Both approaches incorporate instruction in the native language of the non-English-dominant student, while introducing English as a second language. Unlike bilingual/bicultural education programs, however, TBE's terminate language instruction, "once the student is fully functional in the second language." According to the 1975 HEW recommendations, therefore, districts adopting the TBE approach must provide predictive data which show that students are ready to make the transition from Spanish reading into English reading and will succeed educationally in the content areas of the educational programs in which they are to be placed. Thus, if a district elects to accommodate the needs of non-English-dominant students through TBE's, the district must furnish some assurance, based on valid and reliable evidence, that those students will profit meaningfully from their educational program, once native language instruction is terminated.

In a transitional situation, the importance of making reasonably accurate estimates of potential performance cannot be exaggerated. For students for whom English is a second language (ESL), the consequences of premature or delayed transition can be harmful. Students who are transferred too early not only confront inordinate difficulties in trying to keep pace with instruction, but may experience damaging psychological effects as well. These students may be labeled, both by their peers and their teachers, as "slow learners" or "low achievers." As a result, they may be instructionally and/or socially isolated and may never gain instructional parity with their classmates or become fully integrated into the social environment of the school. Needless to say, the situation produced by premature transition also fails to satisfy the fundamental legal mandate of "meaningful instruction."

Students whose transfer or transition is unduly delayed may also suffer unfortunate consequences. Since the ability of most schools to provide a fully comprehensive bilingual curriculum is limited by lack of human and financial resources, students whose transition is delayed may increasingly miss learning opportunities afforded to their peers in the mainstream. Moreover, the ability of many transitional bilingual programs to keep pace with individual instructional needs is often impaired by the necessity to continually accommodate newly-arrived students. The more immediate, and often more serious, needs of these students can understandably overshadow the needs of students who have already mastered fundamental language and other proficiencies. Aside from instructional disadvantages, however, the social and psychological effects of delayed transition may prove harmful as well. When the goal of the bilingual program is transition to the mainstream, students who are held back from transition may begin to perceive themselves as "low achievers," "slow learners," or the like. These impressions, often reinforced by the views of teachers and peers in the mainstream, seem to be further authenticated when such students do enter the mainstream and find themselves unfamiliar with many of the skills that have been acquired by their classmates. Pressures to succeed then become almost overwhelming; those who are motivated may spend the rest of their school years simply trying to "catch up" to their peers.

In response to the judicial mandate of the Supreme Court and to the conspicuous need of schools for materials to aid in instructional transition decisions, NIE initiated a contract with the Center for the Study of Evaluation (CSE) in 1976 to conduct research related to the development of materials to predict post-transition student performance. The initial

stages of the contract were devoted to a pilot study of third grade students in transition from Spanish-English bilingual programs to monolingual English instructional settings. The pilot research utilized a small sample of predominantly native Mexican American students in the greater Los Angeles Metropolitan Area. Results from this study were used in the design of a larger longitudinal study with refined instrumentation and a more broadly based sample. This report describes the initial phase of this two-year study and discusses some of the practical, conceptual, and technical issues involved in the research.

Design of the Study

The extended sample study was designed to explore several research questions. Of primary concern was the identification of variables that might be employed to predict the potential of students to succeed in monolingual English classrooms following transition from bilingual programs. Although consideration was given to various background and demographic factors, the principal focus of the research was on language proficiency variables--specifically reading comprehension, verbal production, and aural comprehension in English and the native language which in this case was Spanish.

Of secondary and collateral concern was the validity and reliability of available instruments that might be used to measure these various language proficiencies. Based on analyses from the pilot research, refinements were made in the initial battery of language proficiency tests. To explore the possibility of test-or-battery-specific results, a second and parallel battery of tests were constructed, based on extensive item analysis of existing, commercially available tests. Because different

tests measure different sub-skills to different degrees, the construction of the test batteries followed a specification of language proficiency skills that had earlier been identified in the pilot research. Each of the batteries included tests of reading comprehension, verbal production, and aural comprehension in English and Spanish. Appendix A provides a list of the objectives or skills that were specified for construction of the two test batteries.

A third concern of the research was the comparability of results between different Hispanic language groups. The pilot research had focused exclusively on students of predominantly native Mexican and Mexican American origin in the Los Angeles area. For the extended sample study, a second Hispanic language group was selected--students of predominantly Puerto Rican and Dominican origin from the East Coast of the U.S. Although there was considerable justification for including yet other Hispanic groups--e.g., Mexican American students in Texas border communities, students of Cuban origin in the Southeast--financial and logistical constraints precluded the involvement of additional sub-populations in this research.

The initial design called for two cohort groups: the current year's third-grade students enrolled in transitional bilingual programs, and the students who had been enrolled in third grade transitional bilingual programs the previous year. The students in the latter cohort group were presumed to have made the transition to fourth grade monolingual English classrooms. Separate analyses of these two cohort groups aimed to explore the relationship between language proficiency and classroom performance both prior to and following transition. Difficulties in data collection

and in the identification of an adequate sample of the previous year's third-graders made the intended analysis of these groups unfeasible, however. These problems are more fully discussed in a latter section of this report.

A second phase of the research design called for a follow-up of the current year's third-graders subsequent to their transition to monolingual English classrooms. This phase of the study is planned for the winter of 1978, with preliminary results available early in 1979.

Selection of the Sample

Target sites were identified on both the East and West Coasts. The East Coast sample was to consist of approximately 200 students of predominantly Puerto Rican and Dominican origin, while the West Coast sample was to consist of approximately 200 students of predominantly native Mexican and Mexican-American origin. On each coast, half of the sample was to be drawn from current third-grade students enrolled in transitional bilingual programs terminating at the end of the third grade; the second half of the sample on each coast was to be drawn from the previous third-grade students enrolled in such programs in the same schools.

The sample was further split on each coast and between grades by test battery. Half of each sample on each coast received Battery I, while half received Battery II.

Difficulties were encountered from the beginning in the identification of target sites, particularly on the East coast. Three major school districts with large Hispanic populations were approached but declined to participate in the study. Ultimately, schools in three smaller school districts agreed to participate. Each district on both the East and West

Coasts was provided with detailed specifications for identifying the sample of students to be tested. The East Coast districts had some difficulty in identifying adequate numbers of the previous year's third-graders who had been enrolled in transitional bilingual programs. The reasons for this were not totally clear. In some cases, districts reported that the students had moved or transferred to other schools in the district; in other cases, it was merely indicated that sufficient numbers did not exist. As a result, however, the size of the cohort group of last year's third-graders was inadequate to conduct the analyses that were originally intended. A description of the size and breakdown of the actual sample is provided in the section of this report titled "Research Results."

Data Collection

Administration of the test batteries to the sample of students was conducted by project staff with the assistance of classroom teachers at the various sites. Each battery was administered to half of the total sample on the East and West coasts. Tests of reading and aural comprehension were group-administered, as was a project-developed test of attitudes toward school. Tests of verbal production were individually administered.

In addition, classroom teachers were asked to complete a Student Information Sheet for each child in the sample. The Student Information Sheet called for general background information, such as age, sex, place of birth, as well as for information regarding language usage and exposure to English. Teachers were also asked to make normative judgments regarding the student's current ability to profit from English language instruction and his/her performance levels across five subject matter areas: reading, language arts, mathematics, science, and social studies. A copy of the Student Information Sheet is provided in Appendix B.

Research Results

Sample

The sample originally consisted of 396 students. Because teachers did not complete the Student Information Sheet for all pupils, student information was available for only 335 children. Of these students, 242 (72.2%) were third graders, and 93 (27.8%) were fourth graders; 182 (55%) were male, and 149 (45%) were female; 178 (53.1%) were from the West coast, and 157 (46.9%) were from the East coast. A breakdown of the numbers by grade, region, sex, and ethnicity is given in Table 1. As can be seen, all the Mexican-American and native Mexican children (46.2% of the sample) were from the West coast, while almost all the Puerto Rican and Dominican pupils (46.2% of the sample) were from the East coast (only one exception). Thus, in this sample ethnicity and region are completely confounded.

Table 2 compares the two regions/ethnicities on several background and instructional variables. Items from the Student Information Sheet served as the dependent variables. Thus items concerning the language which the child used when communicating with parents, peers, and siblings were summed to form a single English usage measure. Comparisons were performed separately for the third and fourth grades. Examining the third grade data first, significant differences between the East and West coasts were found for length of reading instruction in English ($F=10.39$, $df=1/168$, $p=.0015$), months enrolled in school ($F=11.02$, $df=1/181$, $p=.0011$), and English usage with parents, peers, and siblings ($F=4.40$, $df=1/236$, $p=.0370$). Thus, students on the West coast were enrolled longer, had more reading instruction in English, and were more likely to use English when speaking in informal settings.

The results were similar for the fourth grade students. The West received more reading instruction in English ($F=5.05$, $df=1/73$, $p=.0276$) and were more likely to speak English to parents, siblings, and peers ($F=6.38$, $df=1/83$, $p=.0134$). However, fourth graders from the West coast had not been enrolled longer. In addition, the Puerto Rican and Dominican students in the fourth grade had been residents in this country longer ($F=4.89$, $df=1/59$, $p=.0309$) and were six months older ($F=5.09$, $df=1/75$, $p=.0270$).

A breakdown for type of instructional program by grade and region/ethnicity is presented in Table 3. The transition from bilingual to monolingual instruction can be seen by comparing the third and fourth grades. Over 80% of the third graders were receiving bilingual instruction, while less than 3% of the fourth graders fell into this category. Furthermore, less than 5% of the third grade students were in monolingual English programs compared to more than 70% of the fourth graders. Some differences between regions should be noted. The regions differed in the amount of monolingual English and Spanish maintenance instruction given the third graders (11.6% for the West and 2.3% for the East). In addition, the transition to monolingual English instruction in grade four seemed less complete for the West (66.7%) than for the East (79.2%). Thus, the West coast schools placed more emphasis on English instruction in grade three but less emphasis in grade four.

TABLE 1

	GRADE 3				GRADE 4				
	West		East		West		East		
	Male	Female	Male	Female	Male	Female	Male	Female	Total
Mexican-American	24	20	--	--	11	15	--	--	70 (21.1%)
Native Mexican	24	26	--	--	15	18	--	--	83 (25.1%)
Puerto Rican	--	--	73	41	--	1	12	9	136 (41.1%)
Dominican	--	--	9	6	--	--	--	2	17 (5.1%)
Cuban	1	1	--	--	1	--	--	--	3 (.9%)
Central American	2	2	--	1	2	1	1	--	9 (2.7%)
South American	--	--	1	--	--	--	--	--	1 (.3%)
Other	4	2	--	1	2	3	--	--	12 (3.6%)

Grade 3	242 (72.2%)
Grade 4	93 (27.8%)
Male	182 (55%)
Female	149 (45%)
West	178 (53.1%)
East	157 (46.9%)

TABLE 2

	GRADE 3		GRADE 4	
	West	East	West	East
Background Variables:				
Length of residence (months)	49.9	55.9	59.5	84.3
Age (years as of Sept. 1977)	8.5	8.6	9.1	9.6
English Usage	2.8	2.4	3.3	2.5
Instructional Variables:				
Months in Bilingual Class	18.4	17.6	26.1	24.5
Months of reading in English	19.8	13.6	30.5	21.6
Months of reading in Spanish	16.5	16.7	25.2	23.9
Months enrolled	29.1	24.2	37.3	34.8

TABLE 3

	Monolingual English	English & Spanish Maintenance	Bilingual	Spanish & ESL	Monolingual Spanish
Third Grade					
West	9 (8.7%)	3 (2.9%)	83 (80.6%)	8 (7.8%)	---
East	2 (1.5%)	1 (1.8%)	113 (86.9%)	14 (10.8%)	---
Total	11 (4.7%)	4 (1.7%)	196 (84.1%)	22 (9.4%)	---
Fourth Grade					
West	32 (66.7%)	11 (22.9%)	2 (4.2%)	2 (4.2%)	1 (2.1%)
East	19 (79.2%)	1 (4.2%)	---	4 (16.7%)	---
Total	51 (70.8%)	12 (16.7%)	2 (2.8%)	6 (8.3%)	1 (1.4%)

NOTE: Percentages are percent of row

Instruments

The goal of the Bilingual Prediction Study was to predict the fourth grade performance of students who have made a transition from bilingual to monolingual English instruction. The prediction equation was obtained by regressing fourth grade performance on the scores from the third grade test batteries. In order to construct one equation for all students, regardless of region/ethnicity or test pattern, two assumptions had to be made. First, the tests from both batteries must be measuring the same constructs. Second, the tests must be measuring the same constructs for students from both regions. If these assumptions are violated, then separate prediction equations would have to be constructed for each region and each battery. While testing these assumptions of construct invariance across battery and population, the internal-consistency reliability of each test was examined for each region/ethnicity to ensure that the constructs were being measured with the same accuracy.

Internal consistency reliability for each test in both batteries (except the Battery I tests of verbal production in which items were not scored individually) were computed separately for both regions (see Table 4). The coefficients were somewhat lower than would be expected for normative tests. This can be attributed to the large variability in item difficulty usually found in objective-based tests of this type. In particular, the reliability for Battery II, Spanish verbal production was lower than expected. The patterns of reliability were similar for both regions, although there were some differences.

In an effort to determine if the tests were measuring the same constructs in each region/ethnicity, several split-plot ANOVAs were computed. In each ANOVA the factors were region/ethnicity (between-subjects factor)

and the items for that particular test (within-subjects factor). Significant differences between regions were found for three of the six Battery I tests: English listening ($F=13.16$, $df=1/116$, $p<.001$), English reading ($F=37.11$, $df=1/110$, $p<.001$), and Spanish reading ($F=4.33$, $df=1/118$, $p<.040$). Main effects for region were also found for Battery II: Spanish listening ($F=9.79$, $df=1/92$, $df<.002$) and Spanish reading ($F=4.34$, $df=1/92$, $p<.04$). These results would not concern us if the differences were consistent across batteries (i.e., considering the longer English instruction received in the West, higher English language scores would be expected). However, as Table 5 shows, the differences were consistent only for half the tests: Spanish reading, English speaking, and Spanish speaking. The West performed significantly higher on English listening and English reading for Battery I, but there were no differences on the same tests for Battery II. On the other hand, the East had higher Spanish listening scores on Battery II but not on Battery I.

Further evidence regarding construct invariance across regions was provided by the region-by-item interaction term in the ANOVA. Significant interactions (using degrees-of-freedom for the conservate F-test) were found for Spanish listening ($F=4.03$, $df=1/108$, $p<.05$) and Spanish reading ($F=5.39$, $df=1/118$, $p<.05$) on Battery I, but only for Spanish listening on Battery II ($F=9.00$, $df=1/92$, $p<.01$).

As a test of construct invariance, the four intercorrelation matrices reported in Tables 6 and 7 were computed. If both batteries were measuring the same constructs in both populations, then the factor structure for the four combinations of test battery and region-ethnicity should be identical. Because of the small number of subjects within each combination

with complete data, factor analysis was not appropriate. Inspection of the correlation matrices, however, does provide some information.

The six tests supposedly measure three distinct constructs (listening comprehension, verbal production, and reading comprehension). Furthermore, since these constructs are are given in two languages, there should also be two language factors (see Figure 1). This factor structure would produce a correlation matrix in which all the Spanish and English tests would be intercorrelated; but across languages, only tests of the same constructs would be correlated.

Battery I tests will be discussed first. The correlations among the English tests for both regions/ethnicities indicate the existence of an English language factor. However, the magnitude of these correlations are lower for the East than the West. The Spanish reading test also appears to be related to this factor for the West but not the East. A clear Spanish language factor is not evident for either region. The intercorrelations for the East and the West are discrepant enough for one to seriously question whether Battery I tests are measuring the same constructs for the two populations.

The analysis of Battery II leads to similar conclusions. There is an English language factor for both regions, and Spanish reading seems to be related to this factor for the West but not the East. As before, there are enough differences between the two populations to suggest that the same constructs are not being measured.

Comparison of the correlation matrices for each battery indicates some differences in the constructs measured. First, the correlations among the English tests are higher overall for Battery I. Although both batteries have an English language factor, the factor may be more pronounced in

Battery I. Second, the correlation between the English and Spanish reading tests are larger for Battery I than Battery II, indicating a clearer reading comprehension factor for Battery I. Moreover, the Spanish listening comprehension test for Battery I is not correlated significantly with any of the other tests for the West coast students, raising some question as to what the test is actually measuring. Given the number of correlation coefficients calculated, it is difficult to assess the degree of correspondence between the batteries. However, enough differences occur to suggest that there may not be construct invariance across the batteries.

TABLE 4

	# Items	Coefficient alpha West	East
Battery I			
English listening	30	.80	.72
Spanish listening	30	.69	.74
English reading	38	.87	.92
Spanish reading	30	.86	.78
Battery II			
English listening	30	.88	.87
Spanish listening	30	.75	.89
English speaking	14	.70	.70
Spanish speaking	14	.52	.53
English reading	45	.87	.80
Spanish reading	44	.91	.88

TABLE 5

	Battery I		Battery II	
	West	East	West	East
English listening	.589	.492*	.637	.580
Spanish listening	.673	.630	.514	.702*
English reading	.488	.262*	.424	.422
Spanish reading	.586	.517*	.533	.458*
English speaking	3.69	3.37	.649	.613
Spanish speaking	4.72	4.48	.902	.867

TABLE 6

Battery I (West Upper Triangle and East Lower Triangle)

	ELIST	SLIST	ESPEK	SSPEK	EREAD	SREAD
ELIST	--	.19 (50)	.41** (48)	-.14 (52)	.64** (51)	.53** (53)
SLIST	.40** (52)	--	-.11 (44)	-.07 (48)	.18 (48)	.19 (49)
ESPEK	.41** (54)	.09 (51)	--	-.28* (48)	.63** (45)	.38** (48)
SSPEK	.40** (52)	.31* (50)	.27* (58)	--	-.09 (50)	.40** (50)
EREAD	.23* (54)	.09 (54)	.62** (51)	.10 (52)	--	.60** (50)
SREAD	.18 (60)	.44** (51)	.19 (54)	.18 (53)	.44** (54)	--

*p<.05

**p<.01

TABLE 7

Battery II (West Upper Triangle and East Lower Triangle)

	ELIST	SLIST	ESPEK	SSPEK	EREAD	SREAD
ELIST	--	.23 (34)	.33* (42)	.33* (43)	.57** (45)	.10 (35)
SLIST	-.03 (48)	--	-.30* (35)	-.09 (33)	.27 (34)	.32* (36)
ESPEK	.44** (38)	-.06 (43)	--	.20 (44)	.39** (42)	-.18 (36)
SSPEK	.07 (41)	.30* (45)	.21 (42)	--	.21 (42)	.36* (35)
EREAD	.53** (43)	.07 (46)	.41** (36)	-.06 (37)	--	.40** (34)
SREAD	.01 (43)	.14 (50)	.32* (42)	.14 (44)	.08 (42)	--

*p<.05

**p<.01

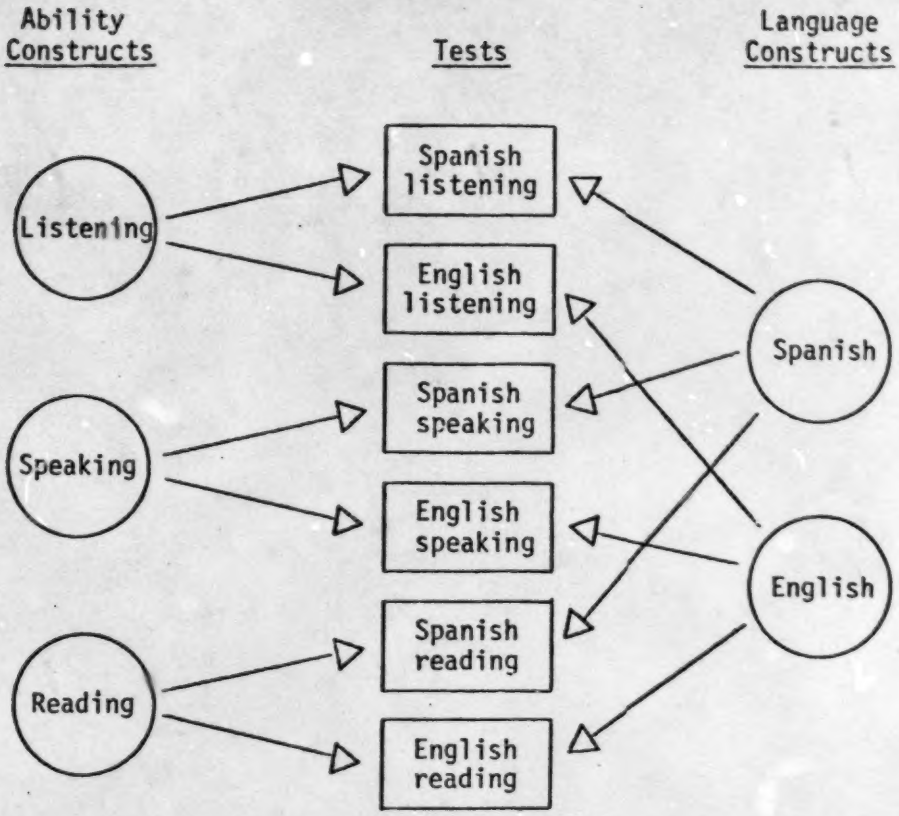
Ability Constructs
Tests
Language Constructs
Listening
Speaking
Reading
Spanish listening
English listening
Spanish speaking
English speaking
Spanish reading
English reading
Spanish
English

Figure 1

Regression Analyses

Two further analyses could be performed given the data available at this stage of the Prediction Study. Teachers assessed each student's performance in reading, language arts, mathematics, social studies, and science. In addition, teachers rated each child's ability to profit from monolingual English instruction. In order to determine how these teacher judgments were related to scores from the standardized test batteries, two sets of regression analyses were performed--one using the performance ratings as the dependent variable and the other using ability to profit from English instruction.

Originally, it was planned that the analyses would be conducted separately for third and fourth graders. Unfortunately, this was not possible. Because the previous analyses suggested that corresponding tests from two batteries may not be measuring the same constructs, the regression analyses were run separately for each battery. Moreover, due to the failure to obtain complete data from all subjects, sufficient numbers of subjects were available only from the third graders. As a result, only the data analysis for the third grade will be reported.

Let us consider the measures of student performance first. Teachers made five ratings, one rating for each of the five subject areas. In addition, teachers were asked to specify the language in which instruction was generally given. There was a high degree of correspondence among both the performance ratings (see Table 8) and the language of instruction (see Table 9). Because of this relationship, performance scores and language of instruction scores were averaged, respectively, over the subject areas. The language of instruction variable can be considered as a continuous measure ranging from one to seven. A high score indicates

that greater amounts of instruction were in English, while a low score indicates that less instruction was given in English (i.e., more instruction in Spanish). Since both language of instruction and region/ethnicity may affect the teachers' ratings, these variables were included in the analysis as independent variables (along with the test scores). Furthermore, region/ethnicity by test score interaction terms were also included in order to test for possible regional differences in the regression coefficients.

The regression analysis assumes that the teacher ratings for all students are independent. This assumption cannot be unconditionally accepted when students are nested within classes (i.e., teachers often rated more than one student from their class). How seriously this violation affected the analysis is unknown. Because of the small number of classes, it was not possible to use the class as the unit of analysis. Nor was it feasible to use deviations from mean teacher rating as the dependent variable. The loss of degrees of freedom in order to estimate each teacher's mean rating would have made any meaningful analysis impossible.

The results of the regression analyses are reported in Table 10 (Battery I) and Table 11 (Battery II). The regression of mean performance rating on the variables listed in Table 10 yielded a multiple R of 0.88 ($F=15.05$, $df=14/63$, $p<.01$) and an R^2 of 0.77 (adjusted $R^2=0.72$). Significant main effects were found for region (the East coast students were rated higher by their teachers), language of instruction (the more English instruction, the higher the rating), and several of the tests (Spanish listening, Spanish speaking, and English reading were positively related to performance). These main effects must be interpreted in light of the intercorrelations among the independent variables (see Table 12).

Substantial correlations exist among all the tests given in English, and between Spanish reading and the English language tests. In particular, the simple Pearson r's between performance and the independent variables should be noted. Although the regression coefficient for region was positive, the correlation with performance was negative (r=-.46). This change in sign may be due to the relationship between language of instruction and region (r=-.72). Because of this confounding, it is difficult to separate the effect of region/ethnicity from language of instruction.

In addition to the main effects, one interaction was significant (region by Spanish verbal production). Separate analyses by region were performed in order to determine the nature of this interaction. Performance was significantly related to Spanish (r=.37) for the East coast but not for the West (r=.003).

The regression analysis for students receiving Battery II resulted in somewhat different findings (see Table 11). The multiple R(0.82) and R^2(0.66) were smaller, although still significantly different from zero (F=5.52, df=14/39, p<.01). The adjusted R^2, however, dropped substantially more than in Battery I to 0.54, possibly because of the smaller sample taking Battery II. There was no main effect for language of instruction. As before, the regression coefficient associated with region/ethnicity was significant (teachers from the East rated their students higher than teachers on the West coast). Spanish reading comprehension was the only test with a significant main effect. The analysis, as with Battery I, must be interpreted in view of the intercorrelations among the independent variables (see Table 12). Spanish listening, Spanish reading, and English reading have significant Pearson r's with performance. Although the

correlation between performance and Spanish listening was higher than the correlation between Spanish reading and performance, the regression coefficient for Spanish listening was not significant. This may have been due to the interaction between region and Spanish listening. Performance and Spanish listening were related for the West (r=.67) but not for the East (r=-.05).

TABLE 8

Correlations Among Performance Ratings

	READ	LANG	MATH	SOC SCI	SCIENCE
READ	--	.89	.71	.75	.75
LANG	.92	--	.70	.72	.76
MATH	.87	.92	--	.72	.72
SOC SCI	.91	.93	.96	--	.92
SCIENCE	.89	.91	.93	.99	--

Battery I (lower triangle)

Battery II (upper triangle

TABLE 9

Correlations Among Language of Instruction

	READ	LANG	MATH	SOC SCI	SCIENCE
READ	--	.86	.76	.64	.56
LANG	.92	--	.83	.61	.49
MATH	.90	.96	--	.81	.61
SOC SCI	.95	.97	.97	--	.64
SCIENCE	.95	.97	.97	1.00	--

Battery I (lower triangle)

Battery II (upper triangle)

TABLE 10

Variable	B	Std Error	F
Region	1.26	.34	13.96**
Lang of Inst.	1.20	.15	63.98**
English List	-.003	.031	0.01
Spanish List	.067	.032	4.42*
Eng Speak	-.10	.095	1.14
Span Speak	.70	.22	10.26**
Eng Read	.065	.023	8.13**
Span Read	-.015	.029	0.28
Region x ELIST	-.10	.064	2.63
Region x SLIST	-.043	.068	0.40
Region x ESPEAK	-.053	.19	0.074
Region x SSPEAK	-.93	.44	4.42*
Region x EREAD	.015	.05	.10
Region x SREAD	-.47	.057	.67

*p<.05

**p<.01

TABLE 11

Variable	b	Std Error	F
Region	.71	0.19	14.34**
Lang of Instruct.	.084	0.13	0.41
English listening	.00016	0.016	--
Spanish listening	.0082	0.012	0.47
English Speaking	-.017	0.028	0.35
Spanish Speaking	.016	0.045	0.12
English Reading	.019	0.016	1.34
Spanish Reading	.039	0.011	11.69**
Region x ELIST	-.043	0.034	1.59
Region x SLIST	-.054	0.024	4.99*
Region x ESPEAK	.024	0.057	0.18
Region x SSPEAK	.038	0.089	0.18
Region x EREAD	.036	0.034	1.11
Region x SREAD	-.0020	0.022	0.008

*p<.05

**p<.01

TABLE 12

	ELIST	SLIST	ESPEAK	SSPEAK	EREAD	SREAD	SCHOOL	LANG	PERF
ELIST	--	.10	.33	.23	.47	.12	-.25	.17	.04
SLIST	.28	--	-.30	-.13	.26	.25	.34	-.51	.59
ESPEAK	.34	-.06	--	.06	.32	-.01	-.16	.38	-.19
SSPEAK	.11	.27	-.11	--	.11	.23	-.35	.15	-.07
EREAD	.56	.15	.50	.04	--	.38	.05	.19	.34
SREAD	.55	.33	.49	.10	.68	--	-.19	-.10	.44
SCHOOL	-.26	-.19	.04	-.23	-.53	-.16	--	-.37	.42
LANG	.17	.28	-.07	.33	.38	.09	-.72	--	-.27
PERF	.29	.36	-.01	.35	.47	.26	-.46	.76	--

Battery I (lower triangle)

Battery II (upper triangle)

The regression analysis of the teachers' rating of ability to profit from monolingual English instruction are reported in Table 13 (Battery I) and Table 14 (Battery II). The multiple R for Battery I was .85 ($F=12.30$, $df=13/64$, $p<.01$), the R^2 was .71, and the adjusted R^2 was .66. Three main effects were significant: Region (East rated students higher), English listening comprehension, and English reading comprehension. The interaction between region and Spanish reading comprehension was also significant. Spanish reading was related to profit rating for the West ($r=.84$) but not the East ($r=.21$). The intercorrelations among the independent variables have already been presented in Table 12. Table 15 shows the Pearson correlation coefficients between each test and the profit rating. All the tests given in English and the Spanish reading test are related to profit.

The Battery II analysis resulted in a multiple of R of .75 ($F=4.04$, $df=13/40$, $p<.01$), an R^2 of .57, and an adjusted R^2 of .43. The only main effect was English reading comprehension, although as shown in Table 15, all the English language tests correlated with profit. As in Battery I, region and Spanish reading interacted. However, in this case Spanish reading was related to profit for the East ($r=.44$) but not the West ($r=.10$). Thus, the teachers' rating of ability to profit from English instruction was dependent on the child's English language skills.

TABLE 13

Variable	b	Std Error	F
Region	1.29	.25	26.47**
Englist	.083	.028	8.55**
Spanlist	.018	.029	.39
English Speaking	.16	.086	3.38
Spanish Speaking	-.29	.20	2.10
English Reading	.063	.021	9.19**
Spanish Reading	.016	.026	.38
Region x ELIST	.058	.059	.99
Region x SLIST	.060	.058	1.08
Region x ESPEAK	.21	.18	1.43
Region x SSPEAK	.39	.39	1.03
Region x EREAD	.007	.042	.03
Region x SREAD	-.11	.052	4.73*

*p <.05

**p <.01

TABLE 14

Variable	b	Std Error	F
Region	.27	.35	.58
Englist	.031	.031	.96
Spanlist	-.021	.022	.97
Engspek	.080	.054	2.18
Spanspek	.045	.087	.27
Engread	.084	.031	7.49**
Spanread	.016	.021	.57
Region x ELIST	-.13	.065	3.73
Region x SLIST	-.080	.047	2.94
Region x ESPEK	.0018	.11	--
Region x SSPEK	.14	.17	.69
Region x EREAD	-.041	.064	.40
Region x SREAD	.11	.043	6.21*

*p<.05

**p<.01

TABLE 15

Variable	Battery I	Battery II
School	.25	-.07
Englist	.54	.39
Spanlist	.10	.06
Engspek	.58	.40
Spanspek	-.12	.11
Engread	.46	.52
Spanread	.53	.21

Discussion and Conclusions

A variety of practical, conceptual, and technical issues were raised in the process of conducting this phase of the research. Some of these issues are reflected in the results of the analyses described in the previous section of this report. Others deserve additional emphasis for their potential significance in affecting future phases of the research and analysis.

The difficulties of data collection are alluded to in the initial section of the report. The selection of appropriate sample groups was complicated by the reluctance of a number of school districts to participate in the testing program. It was apparent, particularly in the East Coast sites, that bilingual programs are surrounded by political controversy and are particularly sensitive to the intrusion of external data collection activities. Despite repeated assurance of confidentiality, administrators and teachers alike expressed grave concern that test results might negatively affect the goals and objectives of the programs. Several districts initially approached for participation declined to cooperate on these grounds. Another district, which agreed to participate, became embroiled in a legal-political battle several weeks before the data collection was conducted; when the CSE project representative arrived to initiate data collection, she was met by hostility and reluctance to permit testing to take place.

School districts displayed difficulties in conforming to the specifications for the sample. In particular, it was problematic for districts to identify adequate numbers of students who had been enrolled in third-grade transitional programs during the previous year. As a result, the

sample size of this cohort group was insufficient to carry out the data analysis originally planned.

Furthermore, as the discussion of the sample in the Results section suggests, the delineation between transitional and maintenance language programs is not observed with any degree of uniformity between schools or districts. Not only do some nominally "transitional" programs appear to be more "maintenance" oriented, but in several instances, students who were designated as being in transitional third-grade bilingual programs appeared to be receiving monolingual English instruction.

Another problem occurred in the collection of background information for the sample. In many cases, neither the classroom teacher nor the school office was able to furnish information on students' ages, birthdates, places of birth, etc. Greater difficulties were encountered in attempting to determine the amount of language instruction in English and Spanish that the students had received. Moreover, many schools were unable to furnish information on the total years of schooling of many of the pupils; school records were incomplete in a large number of cases.

Although efforts were made to construct the two test batteries in such a way that they would be comparable in terms of constructs being tested and time needed for test administration, it is apparent from the data analysis that the comparability of the batteries is somewhat questionable. While tests for factorial invariance could not be performed, comparison of the correlation matrices for the two batteries revealed that subtests in different batteries were not measuring the same construct. Though both batteries contained an English language component, regional and cultural variations affected the relationship among the

Spanish language subtests. Moreover, the batteries appeared to be measuring different constructs for the East and West coast populations. Thus, even within the same battery, the meaning of the subtests depended on where the battery was given.

Equally important is the issue of collinearity between the predictor variables of language proficiency. The subtests were chosen because they supposedly tapped district language abilities in both Spanish and English. The intention was to construct profiles of language skills. However, the substantial intercorrelations among the English language subtests prevented the separate consideration of the subtest scores. This dependency also made it difficult to interpret the coefficients from the regression analyses. Two measures of the same construct may both be highly predictive of the criterion. Yet, when both variables are included in the regression equation, neither variable may appear significantly predictive.

Another issue of significance is the type of test instruments that are best utilized in a prediction context. Standardized, norm-referenced tests tend, in general, to be more appropriate for prediction purposes. Few of such instruments have been normed for non-English-speaking minorities, however. Moreover, for those that have, normative data tend to correspond to particular ethnic or regional groups; from the analyses presented in this report, there is some question as to whether the same norms may be applied to other ethnic or regional groups sharing the same native language. Arguments against the use of standardized, norm-referenced instruments include the charge that such tests bear only random or coincidental relationship to instruction or curricular content. Thus, opponents insist, fairer measures consist of criterion-referenced

assessment instruments that are specifically geared to pre-determined sets of proficiencies for which instruction is provided. The two batteries employed in this study were constructed from commercially-available instruments which had been previously normed among certain regional or ethnic groups and for which certain language skills or proficiencies that are normally included in instruction could be specified. By splitting the sample for each battery between the two ethnic/regional groups tested in this study, it was hoped that instrument biases related to ethnicity or region would be identifiable.

Some of the biases that do appear may relate to linguistic differences between the two populations. For example, the initial consonant sound of "j" (as in "jello") is unfamiliar to those of Mexican origin, while it is used by those of Puerto Rican origin in pronouncing an initial "y" consonant sound (as in "yam"). Other differences relate to the labiodeutal "v" and the bilabial "b" sounds; children of Puerto Rican origin, for example, never use the labiodeutal "v" sound. Children of Mexican-American origin, particularly those of Northern Mexican heritage, do not distinguish interconsonantal /o/, /u/, /e/, or /i/. Thus, words like "porito" and "purito" sound identical.

Familiarity with certain vocabulary terms in English and Spanish characteristically differentiates the two populations studied as well. English words such as "basement" and "furnace" are less familiar to West Coast children than to East Coast children. Similarly, the word "pavo" is more characteristic of Caribbean Spanish; Mexican-Americans would substitute "guajolate" or "cocono."

While these and other items in the test batteries represent potential problem areas for the two populations, it is not clear to precisely what extent such differences affected the outcome of the data analysis. Despite efforts to minimize the existence of these potentially troublesome items, both of the test batteries contain some such items. Considerable additional research would be required to explore the effects of these differences further.

For this phase of the analysis, two types of measures were utilized as outcome or dependent variables in the prediction model. The first measure was a composite index of teacher ratings of performance across five subject matter areas (PERFORMANCE). The second measure was a separate teacher judgment index based on the teacher's rating of a student's ability to currently profit from English language instruction (PROFIT). As reported earlier, the regression analyses run on these two variables by separate battery produced dissimilar results, which are instructive in terms of illuminating differences between ethnic/regional groups and between test batteries, but which fail to identify a uniform set of predictor variables. With the follow-up phase of the research, planned for the 1978-79 school year, post-transition performance measures will be collected. These will serve as more accurate dependent measures reflecting student success following actual transition from bilingual settings.

Based on the analysis conducted to date, however, it is possible to draw some tentative conclusions. First, where specific test instruments are to be recommended for use in prediction contexts, the tests should be locally normed. Since what is being predicted is potential for success

in a particular regional setting, measures should reflect normative levels of success within that setting. Second, the relationship between the various language proficiency factors needs to be more fully explored. Although there appears to be an English language factor for both batteries and both ethnic/regional groups, and a Spanish reading factor appears to relate to this in some way, the data need considerably more scrutiny. Because there did not seem to be construct invariance across the batteries, the test battery results were not combined in any of the analyses. A possible avenue of exploration might be to administer the two batteries--or corresponding tests from each battery--to a smaller sample, i.e., administer both Spanish listening tests to a sample of students, both English listening tests to another comparable sample, etc. (Although the ideal would be to administer all tests to the same sample, this is unrealistic in terms of time and unreasonable in terms of the potentially harmful effects on the children.) While such a procedure would not answer all questions, it might prove to be a fruitful starting point for unraveling some of the difficulties in comparing the two batteries.

Appendix A

BILINGUAL PREDICTION STUDY

LIST OF OBJECTIVES AND SOURCES OF ITEMS

FOR 1977 TEST BATTERY

Reading Proficiency in Spanish and English

WORD ATTACK

Given a word orally that has a single initial consonant and given a set of written words, the learner will identify the written word with the same initial consonant.

Given a written consonant and a set of illustrations of animals or things, the learner will identify the illustration whose name contains the consonant in medial position.

Given a written consonant and a set of illustrations of animals or things, the learner will identify the illustration whose name has the consonant in final position.

Given a set of written words, the learner will identify the word that contains an initial consonant blend.

Given a set of written words, the learner will identify the word that contains a medial consonant blend.

Given a word orally that has a final consonant blend and given a set of written words, the learner will identify the written word with the same final consonant blend.

Given a written vowel and a set of illustrations of animals or things, the learner will identify the illustration whose name contains the vowel in the medial position.

Given a written word and a set of numerals, the learner will identify the numeral that represents the number of syllables that the word contains.

SYNTAX

Given an incomplete written sentence and a set of written words, the learner will identify the form of the word that is appropriate to the context of the sentence.

Given a set of written words, the learner will identify the one with a prefix.

Given a set of written words, the learner will identify the one with a suffix.

Given a set of written sentences of which all are correct but one, the learner will identify the sentence in which the inflected verb form is incorrect.

Given a set of written sentences with the verb missing and a set of written inflected forms of a verb, the learner will identify the form that correctly completes the sentence.

VOCABULARY

Given a written word and a set of illustrations, the learner will identify the illustration of the word.

Given a written word and a set of written definitions, the learner will identify the definition of the word.

Given a written word that has an antonym, the learner will identify the antonym from a set of written words.
Given a written word that has a synonym, the learner will identify the synonym from a set of written words.

LITERAL COMPREHENSION

Given an incomplete written sentence and a set of written words, the learner will identify the word whose meaning is appropriate to the context of the sentence.
Given a written selection and a set of illustrations, the learner will identify the illustration that best expresses the main idea of the selection.
Given a written selection and a set of written sentences, the learner will identify the sentence that best expresses the main idea of the selection.
Given a written selection and a set of written sentences, the learner will identify the sentence that best describes a significant detail of the selection.
Given a written selection from which the main idea can be inferred and given a set of written sentences, the learner will identify the sentence that best expresses the main idea of the selection.
Given a written selection and a set of sentences of the main events of the selection, the learner will sequence the sentences.

Listening Comprehension in Spanish and English

ORAL VOCABULARY

Given a word orally and given a set of illustrations, the learner will identify the illustration of the word.
Given a definition of a word orally and given a set of illustrations, the learner will identify the illustration of the word defined.
Given a definition of a word orally and given a set of written words, the learner will identify the word that corresponds to the definition.

COMPREHENSION

Given an oral selection and a set of illustrations, the learner will identify the illustration that best represents the main idea of the selection.
Given an oral selection and an oral question about one of the details of the selection, the learner will identify from a set of illustrations the one that best answers the question.

Verbal Production in Spanish and English

Given a series of pictures, the learner will respond with appropriate grammatical structure to oral questions about the pictures.

APPENDIX B

Bilingual Prediction Study
Student Information Sheet

This sheet is to be completed by the classroom teacher, who may need to consult school records for some of the information (e.g., number of days absent, years in bilingual program, date, and place of birth). Other information (such as language usually spoken with parents, siblings, peers) may be solicited from the student if the teacher is not certain. In some instances, completely accurate information may not be available; teachers should make reasonable estimates, based on the best information they are able to obtain.

1. Name of student_____

2. School_____ 3. Sex_____

4. Present grade_____ 5. Previous grade (1976-77)_____

6. Date of birth_____ 7. Place of birth_____

8. Ethno-cultural membership (circle one):

   1. Mexican-American
   2. Native Mexican
   3. Puerto Rican
   4. Dominican
   5. Cuban
   6. Other Caribbean
   7. Other Central American
   8. South American
   9. Other (specify):_____

9. Years of residence in continental U.S._____

For the following four questions, consider one school year equivalent to nine months.

10. Months enrolled in a bilingual class_____

11. Months of reading instruction in English_____

12. Months of reading instruction in Spanish_____

13. Total months enrolled in (any) school_____

14. Type of instruction presently received (check as applicable):

___ 1) Monolingual English ___ 2) English plus Spanish maintenance

___ 3) Bilingual (Eng./Span.) ___ 4) Spanish plus ESL pull-out

___ 5) Monolingual Spanish

Language usually spoken: (check one)

15. With parents ___ 1) Spanish ___ 2) Span./Eng. ___ 3) English

16. With peers ___ 1) Spanish ___ 2) Span./Eng. ___ 3) English

17. With siblings ___ 1) Spanish ___ 2) Span./Eng. ___ 3) English

18. Number of school days completed this year by your school_____

19. Number of days this child has been absent this year_____

20. How do you rate this child's present ability to profit from monolingual English instruction? (check one)

___ 1) very poor ___ 2) poor ___ 3) fair ___ 4) good ___ 5) very good

On a scale from one to five (1=far below average; 5=far above average), please assess this child's performance in each subject area, in comparison to the typical child at grade-level in your school. In addition, please indicate the language in which instruction is generally given to this child (1=Spanish; 2=English/Spanish; 3=English).

	Subject Area	Level of Performance					Language of Instruction		
21.	Reading	1	2	3	4	5	1	2	3
22.	Language Arts	1	2	3	4	5	1	2	3
23.	Mathematics	1	2	3	4	5	1	2	3
24.	Social Studies	1	2	3	4	5	1	2	3
25.	Science	1	2	3	4	5	1	2	3